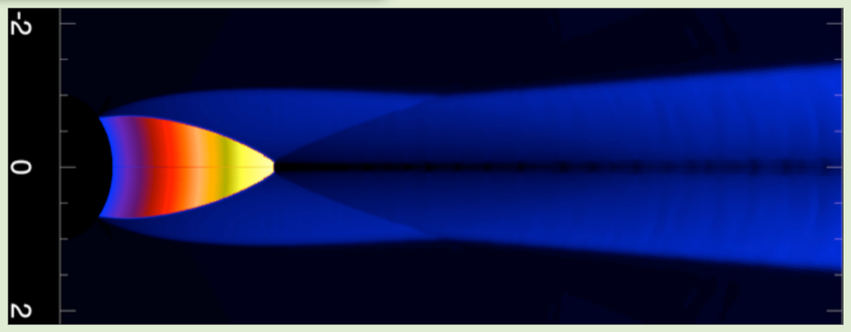
MRGENESIS PERFORMANCE

Carmen Aloy, Petar Mimica & Miguel A Aloy.
Department of Astronomy and Astrophysics. University of Valencia (Spain)

The code MRGENESIS employs a finite volume approach in order to evolve the Relativistic Euler equations in presence of dynamically relevant magnetic fields. This is combined with a Constrained Transport scheme to account for the divergence free evolution of the magnetic field. It incorporates several approximate Riemann Solvers, inter-cell reconstruction techniques and a method of lines based on a Total Variation Diminishing technique to provide up to third order accuracy both in time and in space. The code can also account for the evolution of the non-thermal particles using finite difference methods which are implicit or semi-implicit.

MRGENESIS is implemented in Fortran 90 and parallelized with a hybrid MPI/OpenMP model. Parallel I/O is obtained by using the Hierarchical Data Format (HDF5).



1) VALIDATION TEST

Tests to check that the parallel code obtains the same results as sequential code:

- Local comparison: confront the **analytical** and the **numerical** solution of selected test beds.
- Global comparison: Calculating the integral and getting the **relative error** between analytical and numerical solutions.
- Computing the **order of the method** from the numerical results at in progressively finer grids.

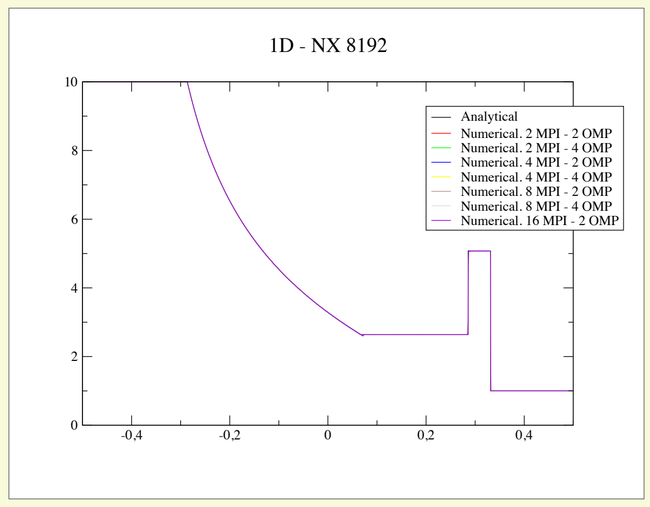
HYBRID CODE MPI-OMP: CALCULATE RELATIVE ERROR BETWEEN ANALYITIC AND NUMERICAL SOLUTION FOR NX = 1000
Analytic integral: 4.9377868599999255

Numerical integral (2 MPI x 4 OpenMP): 4.9376604109999951
RELATIVE ERROR: **2.370E-002**

Numerical integral (4 MPI x 2 OpenMP): 4.9376584239999932
RELATIVE ERROR: **2.376E-002**
Variations among different partitions (OpenMP-MPI) originate from the differences in the roundoff error accumulation.

HYBRID CODE MPI-OMP: CALCULATE THE ORDER OF THE METHOD
EXPECTED ORDER 1 (due to the presence of shocks in the solution)
Relative error (2 MPI x 4 OpenMP, 1000 points): 2.370E-002
Relative error (2 MPI x 4 OpenMP, 2000 points): 1.247E-002
ORDER: **0.92691682464452296**

Relative error (2 MPI x 4 OpenMP, 1000 points): 2.376E-002
Relative error (2 MPI x 4 OpenMP, 2000 points): 1.250E-002
ORDER: **0.92632524106233094**



2) SCALABILITY TESTS ON MARE NOSTRUM (BSC)

MareNostrum has 31 racks with 6 Blade Centers per rack. A Blade Center has 14 Server Blade JS21, which has 2 dual-core processors PowerPC 970MP@2.3 GHz with 1 MB of L2 cache. Blade nodes JS21 are interconnected through Myrinet.

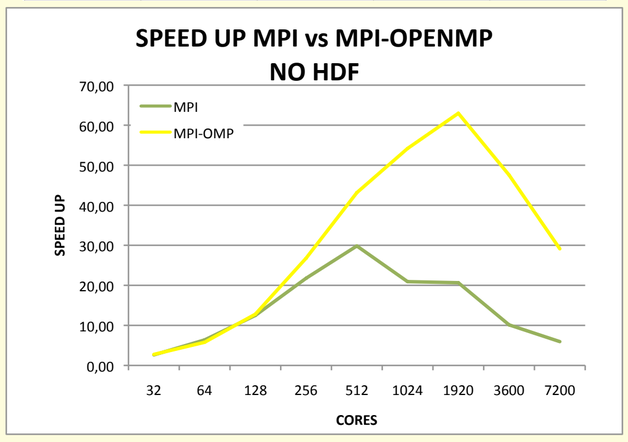
STRONG SCALING TESTS:
We run a standard hydrodynamic problem (the interaction of stellar winds) with different numbers of processors and combinations of MPI and OpenMP threads.

TEST SET UP (MEDIUM SIZED PROBLEM)
GRID: 2880 x 2880
Compiler: IBM XLF90
Optimization options: -O3 -qstrict -q64 -qtune=ppc970 -qarch=ppc970 -qcache=auto
Iterations: 200
SpeedUp relative to 16 CPUs.

2880x2880	WIND	maxiter: 200		
CPUS	MPI	OPENMP	TIME	SPEEDUP
16	16	1	1.611,00	
32	32	1	614,00	2,62
64	64	1	253,00	6,37
128	128	1	129,00	12,49
256	256	1	74,00	21,77
512	512	1	54,00	29,83
1024	1024	1	77,00	20,92
1920	1920	1	78,00	20,65
3600	3600	1	159,00	10,13
7200	7200	1	271,00	5,94

2880x2880	WIND	maxiter: 200		
CPUS	MPI	OPENMP	TIME	SPEEDUP
16	4	4	2.330,00	
32	8	4	847,00	2,75
64	16	4	403,00	5,78
128	32	4	182,00	12,80
256	64	4	87,00	26,78
512	128	4	54,00	43,15
1024	256	4	43,00	54,19
1920	480	4	37,00	62,97
3600	900	4	49,00	47,55
7200	1800	4	80,00	29,13

RESULTS
- The pure MPI model runs faster because the latency introduced by the management of OpenMP threads is greater than the latency for MPI on a small number of processors.
- However, *when the number of processors increases*, the message passing latency grows and penalizes significantly the execution time compared to the hybrid model.
- The hybrid model saturated with more processors, 1920 processors, while the pure MPI model saturated with 512.

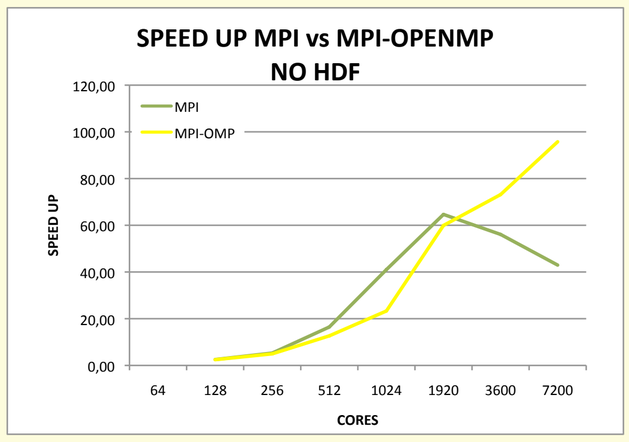


TEST SET UP (LARGE PROBLEM)
GRID: 8640 x 8640
Compiler: IBM XLF90
Optimization options: -O3 -qstrict -q64 -qtune=ppc970 -qarch=ppc970 -qcache=auto
Iterations: 600
SpeedUp relative to 64 CPUs.

8640x8640	WIND	maxiter: 600		
CPUS	MPI	OPENMP	TIME	SPEEDUP
64	64	1	19.071,00	
128	128	1	7.563,00	2,52
256	256	1	3.608,00	5,29
512	512	1	1.159,00	16,45
1024	1024	1	465,00	41,01
1920	1920	1	295,00	64,65
3600	3600	1	340,00	56,09
7200	7200	1	444,00	42,95

8640x8640	WIND	maxiter: 600		
CPUS	MPI	OPENMP	TIME	SPEEDUP
64	16	4	18.577,00	
128	32	4	7.563,00	2,46
256	64	4	3.752,00	4,95
512	128	4	1.468,00	12,65
1024	256	4	797,00	23,31
1920	480	4	310,00	59,93
3600	900	4	254,00	73,14
7200	1800	4	194,00	95,76

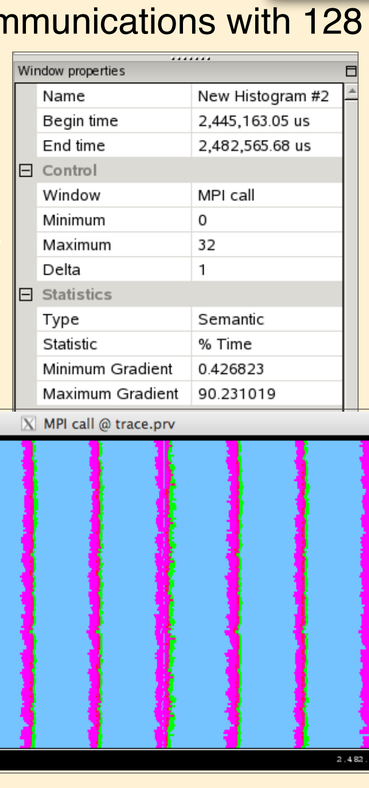
RESULTS
- The mesh size has a significant impact on scalability.
- A larger mesh size increases the work-per-core, reduces the ratio number of numerical zones to be communicated across domains to number of zones in each domain.
- The **hybrid code keeps scaling (though sub-linearly) up to 7200 processors**.

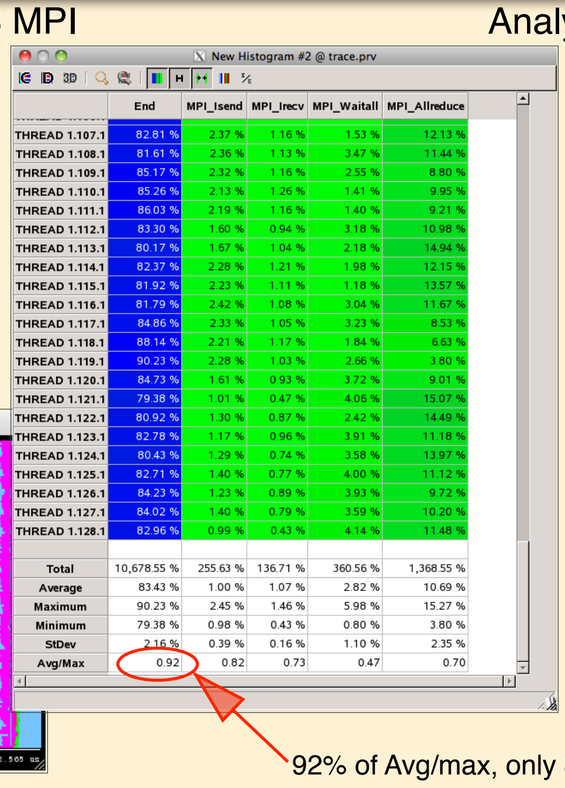


3) ANALYSIS WITH THE HPC TOOL PARAVER

Analysis of communications with 128 MPI

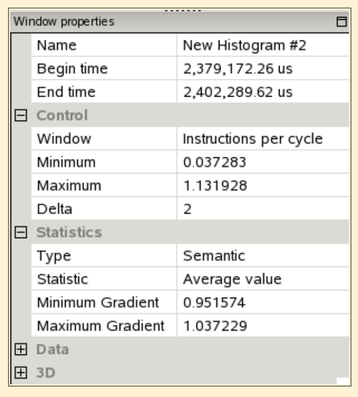
% Time in MPI Calls vs % time doing operations





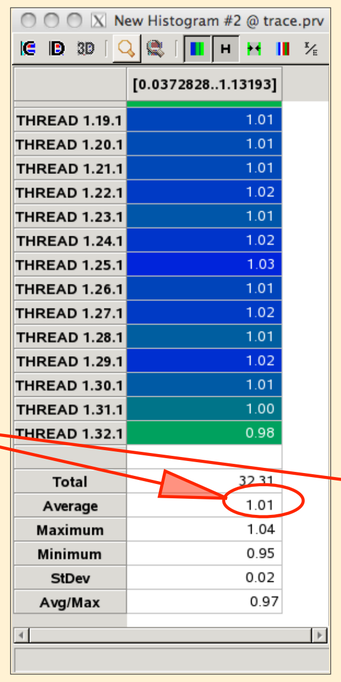
92% of Avg/max, only 8% of time in communications

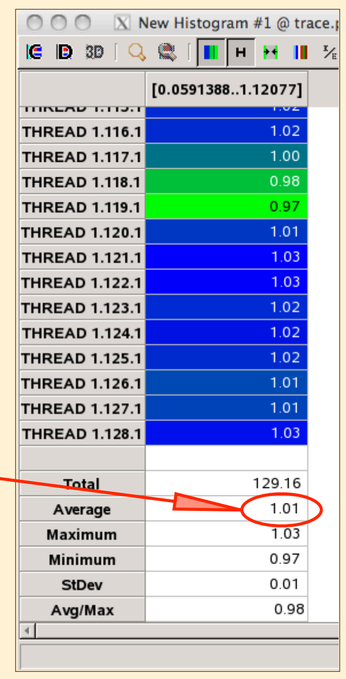
Analysis of communications, IPC comparison between 128 MPI and 32 MPI



Average of IPC is 1.01 in both executions.

The code efficiency is maintained even increasing the number of threads





References

1. "Deceleration of arbitrarily magnetized gamma-ray burst ejecta: the complete evolution"
Mimica. P, Giannios D., Aloy M.A.
Astronomy and Astrophysics 494, 879 (2009)
http://bit.ly/nP3rMz

2. "Internal shocks in relativistic outflows: collisions of relativistic shells"
Mimica P., Aloy M.A., Müller E.
Astronomy and Astrophysics 466, 93 (2007)
http://bit.ly/oVB4Vc

3. "Synthetic X-ray light curves of BL Lacs from relativistic hydrodynamic simulations"
Mimica P., Aloy M.A., Müller E., Brinkmann W.
Astronomy and Astrophysics 418, 947 (2004)
http://bit.ly/qkPRij